

# MMPI における faking の逆説効果と項目内容の異常性判断との関係

## — D 尺度について —

木 村 敦 子

(金沢大学 文学部)

木村(1987a)は MMPI の抑うつ尺度項目について判断者に true と false のどちらの回答が抑うつ採点キーだと思ふかの判断を求め、その適中率を項目の subtlety の指標とした。適中率50%以上の項目を obvious 項目、適中率50%未満の項目を subtle 項目と分類したところ obvious 項目47個、subtle 項目13個に分かれた。

被検者に、意図的に偏った受験態度で検査を受けるように求める faking 教示を用いて MMPI (抑うつ尺度項目)を実施したところ、subtle 項目においては、被検者は人から良い印象を持たれるような方向に回答しようとする (fake good)、意図に反して却って抑うつ尺度に採点される側へ回答し、人から悪い印象を持たれるような方向に回答しようとする (fake bad)、逆に抑うつ尺度に採点されない側へ回答した (木村, 1987b)。この現象を Burkhart, Christian, and Gynther (1978) に倣い「逆説効果」と呼ぶことにする。この逆説効果は obvious 項目では生起せず、被検者は意図に添う方向へ、即ち fake good 条件においては抑うつ尺度に採点されない側へ、fake bad 条件においては抑うつ尺度に採点される側へ回答した。木村(1987b)は、subtle 項目で生じる逆説効果と、obvious 項目で生じる、被検者の意図に添う方向への回答の変化を併せて subtlety 現象と呼んだ。

ところで、抑うつ尺度は精神的な異常性を表わす尺度の一つである。subtlety 現象と同様の現象は、例えば「この項目にどう回答すると精神的におかしいと考えられるか」といった、一般的な精神的異常性の指標を用いた時にも見られるであろうか。本研究ではこの点について確かめてみたい。即ち、「採点キー側へ回答することが精神的におかしいと判断される項目に対しては、被検者は faking 教示において意図に添う方向へ回答を操作できるが、採点キー側へ回答することがおかしいと判断されない項目に対しては、被検者は意図とは逆の方向へ

回答する」という現象が見られるかどうかを調べる。なお、この現象をここでは仮に「reversal 現象」と呼んでおく。reversal 現象は木村(1987b)の subtlety 現象とはほぼ等しいが、木村(1987b)の subtlety 指標よりも一般的な、精神的異常性についての指標で見出される現象を指すことにする。従って、逆説効果は reversal 現象の一部である。即ち、「採点キー側へ回答することがおかしいと判断されない項目については被検者は意図とは逆の方向へ回答する」ことが逆説効果である。

項目の一般的な精神的異常性を表わす指標として二つの指標を用いた。一つは true と false のどちらに回答すれば「より精神的におかしい・問題であるか」の判断 (以後これを「項目の問題側回答方向の判断」と呼ぶ) に基づく指標である。第Ⅰ部では、この指標を用いて、reversal 現象の有無を調べる。もう一つの指標は、項目の抑うつ採点キー側に答えることが「どれくらい精神的におかしい・異常であるか」の評定 (以後これを「項目の回答に対する異常性の評定」と呼ぶ) に基づく指標である。第Ⅱ部ではこの指標を用いて reversal 現象の有無を調べる。

### 第Ⅰ部 項目の問題側回答方向の判断と reversal 現象

抑うつ尺度項目について、「はい」、「いいえ」のどちらに答えると「精神的におかしい・問題がある」ことになると考えられるかの判断を被検者に求め、その結果を精神的異常性の指標とした場合に reversal 現象が生起するかどうかを確かめた。まず項目を、抑うつ採点キー側回答を問題ありと判断する人が一定基準より多い項目 (以後これらの項目を便宜的に「D (AI) 項目」と記す) と、少ない項目 (同じく「D (NI) 項目」と記す) に2分した。次に標準手続き及び faking 教示の下での被検者の抑うつ得点を、D (AI) 項目群の得点と D (NI) 項目群の得点に分けて算出した。fake good 条件におい

ては、D (AI)項目群の得点が標準手続きよりも減少し、D (NI)項目群の得点が上昇すれば reversal 現象が生起したといえる。また fake bad 条件においては、D (AI)項目群の得点が上昇し、D (NI)項目群の得点が減少すれば reversal 現象が生起したといえる。この分析においてはD (AI)項目とD (NI)項目の分割点のとり方を2種類用いた。そのほか、抑うつ採点キー側を問題であるとした判断の各項目における出現率を精神的異常性の指標として、これと標準教示及び faking 教示下における各項目の抑うつ採点キー側回答の出現率との相関関係についても調べた。

## 方 法

### 被検者と手続き

MMPI金沢大学版(多田, 1959: 田中, 1964)を用いた。「項目の問題側回答方向の判断」は、木村(1987a)のデータを用いた。また、faking (fake good 及び fake bad) 教示での被検者の回答は木村(1987b)のデータを用いた。被検者と手続きの概要は以下に示すとおりであった。

**問題側回答方向の判断:** 看護学校の1年生女子78名を判断者とした。判断者に、抑うつ尺度60項目のそれぞれについて「はい」、「いいえ」のどちらに答えたほうが「精神的におかしい・問題がある」かの判断を求めた。

**faking:** 短期大学及び4年生大学の女子学生245名を被検者とした。このうち120名を fake good 条件、125名を fake bad 条件に割り当てた。被検者は全員、先ず最初に標準教示手続きで抑うつ尺度項目に回答し、そのあと faking 教示に従って回答した。fake good 条件の教示は「他人からできるだけ良い印象を持たれるように答えてください。実際の自分がどうであるかにかかわらず、非常に好ましい性質を持っているように見せかけてもらいたいのです」であり、fake bad 条件の教示は「他人からできるだけ悪い印象を持たれるように答えてください。実際の自分がどうであるかにかかわらず、非常に好ましくない性質を持っているように見せかけてもらいたいのです」であった。fake good 条件に割り当てられた被検者群を FG 群(この群は標準教示条件と fake good 教示条件について回答した)と呼び、fake bad 条件に割り当てられた被検者群を FB 群(この群は標準教示条件と fake bad 教示条件について回答した)と呼ぶ。

## 結 果 と 考 察

問題側回答方向の判断において、抑うつ採点キー側を問題であるとした判断の出現率を以後便宜的に「問題率」と呼ぶ。

と呼ぶ。

### 1. 分割点を問題率50%とした場合の reversal 現象の生起について

問題率が50%以上の項目をD (AI)項目、50%未満の項目をD (NI)項目と見なした。前者をD (AI<sub>50</sub>)項目、後者をD (NI<sub>50</sub>)項目と記す。抑うつ尺度60項目中D (AI<sub>50</sub>)項目は38個、D (NI<sub>50</sub>)項目は22個であった。D (NI<sub>50</sub>)項目群とD (AI<sub>50</sub>)項目群内における抑うつ得点を表1と表2に示した。表1は FG 群についての結果であり、表2は FB 群についての結果である。FG 群においては、D (AI<sub>50</sub>)項目群の得点も

表1. fake good 条件群におけるD (AI<sub>50</sub>)項目群及びD (NI<sub>50</sub>)項目群の抑うつ得点

		D (AI <sub>50</sub> ) 項目群	D (NI <sub>50</sub> ) 項目群	合計得点
標準手続き	M	15.7	15.6	31.2(T=52)
	SD	4.97	2.33	5.67
fake good	M	6.7	13.7	20.5(T=34)
	SD	3.04	2.50	4.52

(N=120)

問題率50%以上の38項目をD (AI<sub>50</sub>)項目、50%未満の22項目をD (NI<sub>50</sub>)項目とした場合の、D (AI<sub>50</sub>)項目群及びD (NI<sub>50</sub>)項目群の抑うつ得点を示した。なお、( )内のTの値は抑うつ尺度のT得点を示す。

表2. fake bad 条件群におけるD (AI<sub>50</sub>)項目群及びD (NI<sub>50</sub>)項目群の抑うつ得点

		D (AI <sub>50</sub> ) 項目群	D (NI <sub>50</sub> ) 項目群	合計得点
標準手続き	M	15.6	15.7	31.3(T=52)
	SD	4.59	2.12	5.33
fake bad	M	28.8	7.1	35.9(T=61)
	SD	6.10	2.60	6.48

(N=125)

問題率50%以上の38項目をD (AI<sub>50</sub>)項目、50%未満の22項目をD (NI<sub>50</sub>)項目とした場合の、D (AI<sub>50</sub>)項目群及びD (NI<sub>50</sub>)項目群の抑うつ得点を示した。なお、( )内のTの値は抑うつ尺度のT得点を示す。

D (NI<sub>50</sub>)項目群の得点も fake good 条件の方が標準条件よりも低かった(D (AI<sub>50</sub>)項目群;  $t=18.561$ ,  $df=119$ ,  $p<0.001$  : D (NI<sub>50</sub>)項目群;  $t=7.147$ ,  $df=119$ ,  $p<0.001$ )。人から良い印象を持たれようとして回答すると、D (AI<sub>50</sub>)項目もD (NI<sub>50</sub>)項目も同じように抑うつ得点が低くなる方向へ回答が変化したのである。即ち、reversal 現象は生起しておらず、それ

はD (NI<sub>50</sub>) 項目において逆説効果が生じていないためである。FB 群においては、D (AI<sub>50</sub>) 項目群の得点は標準条件よりも fake bad 条件のほうが高い ( $t=-19.110$ ,  $df=124$ ,  $p<0.001$ ) が、D (NI<sub>50</sub>) 項目群の得点は標準条件よりも fake bad 条件のほうが低かった ( $t=26.218$ ,  $df=124$ ,  $p<0.001$ )。人から悪い印象を持たれようとして回答すると、D (AI<sub>50</sub>) 項目は抑うつ得点が高くなる方向に回答が変化した。D (NI<sub>50</sub>) 項目は抑うつ得点が低くなる方向に変化した訳であり、reversal 現象が認められた。

## 2. 分割点を問題率30%とした場合の reversal 現象の生起について

問題率が30%未満の13項目のみを D (NI) 項目と見なし、それ以外の47項目を D (AI) 項目と見なした。前者をD (NI<sub>30</sub>) 項目、後者をD (AI<sub>30</sub>) 項目と記す。D (NI<sub>30</sub>) 項目群、D (AI<sub>30</sub>) 項目群内における抑うつ得点は表3に示すとおりであった。FG 群において

表3. fake good 条件群と fake bad 条件群における D (AI<sub>30</sub>) 項目群及びD (NI<sub>30</sub>) 項目群の抑うつ得点

		fake good 条件群 (N=120)		fake bad 条件群 (N=125)	
		D (AI <sub>30</sub> )	D (NI <sub>30</sub> )	D (AI <sub>30</sub> )	D (NI <sub>30</sub> )
標準手続き	M	21.4	9.9	21.3	10.0
	SD	5.36	1.86	5.06	1.67
faking	M	10.6	9.9	32.6	3.3
	SD	3.77	1.64	6.62	2.05

問題率30%以上の47項目をD (AI<sub>30</sub>)項目、30%未満の13項目をD (NI<sub>30</sub>)項目とした場合、各条件におけるD (AI<sub>30</sub>)項目群及びD (NI<sub>30</sub>)項目群の抑うつ得点を示した。

は、D (AI<sub>30</sub>) 項目群の得点はD (AI<sub>50</sub>) 項目と同様に標準条件よりも fake good 条件のほうが低く ( $t=19.687$ ,  $df=119$ ,  $p<0.001$ )、一方D (NI<sub>30</sub>) 項目群の得点はどちらの方向への変化も見られなかった ( $t=0.0$ )。即ち reversal 現象は生起しておらず、それは先と同様D (NI<sub>30</sub>) 項目において逆説効果が生じていないことによる。FB 群においては、D (AI<sub>30</sub>) 項目群の得点は標準条件よりも fake bad 条件のほうが高く ( $t=-14.845$ ,  $df=124$ ,  $p<0.001$ )、D (NI<sub>30</sub>) 項目群の得点は fake bad 条件のほうが低かった ( $t=26.012$ ,  $df=124$ ,  $p<0.001$ )。即ち reversal 現象が生起していた。

D (AI) 項目、D (NI) 項目の分割点のとり方にかかわらず、FB 群では reversal 現象が生じたが FG 群では生じなかった。しかし分割点を問題率30%にとった場合は FG 群においても、D (NI<sub>30</sub>) 項目に対する回答の方向はD (AI<sub>30</sub>) 項目とは異なっており、

少なくとも被検者の意図に添う方向へは変化しなかった。

reversal 現象が生じなかった FG 群においても、D (AI) 項目に対する回答は分割点のとり方にかかわらず一貫している。他人から良い印象を持たれようとして回答すると抑うつ得点が下がるのである。これは reversal 現象のうち半分は成り立っていることを示す。故に、reversal 現象の生起は後の半分、即ちD (NI) 項目で逆説効果が生じるか否かにかかっている。D (NI) 項目に対する回答は分割点のとり方によって異なっており、問題率50%を分割点とした場合にはD (AI<sub>50</sub>) 項目と同じ方向に変化した。30%とした場合には変化が生じなかった。問題率30~50%の項目に対する回答はD (AI) 項目と同方向へ変化する人が多いのであろう。

## 3. 問題率とD側回答率の相関について

項目の問題率と各教示条件における抑うつ採点キー側回答出現率 (これをD側回答率と呼ぶ) との相関を求めた結果が表4である。全ての条件において0.7~0.8と

表4. 項目の問題率と各教示条件における抑うつ採点キー側回答出現率の相関

	fake good 条件群	fake bad 条件群
標準手続き	-0.765	-0.763
faking	-0.828	-0.778

(N=60)

回答者は fake good 条件群120名、fake bad 条件群125名であった。

いう相関が見られた。ただし標準条件及び fake good 条件では負の相関 (無相関検定の結果は FG 群の標準条件 ;  $t=-9.046$ ,  $df=58$ ,  $p<0.001$  ; FB 群の標準条件 ;  $t=-8.990$ ,  $df=58$ ,  $p<0.001$  ; fake good 条件 ;  $t=-11.246$ ,  $df=58$ ,  $p<0.001$ )、fake bad 条件では正の相関 (無相関検定 ;  $t=9.431$ ,  $df=58$ ,  $p<0.001$ ) がある。標準条件及び fake good 条件では、問題率の高い (低い) 項目ほど採点キー側に回答する人が少ない (多い) のに対し、fake bad 条件では逆に、問題率の高い (低い) 項目ほど採点キー側に回答する人が多い (少ない) のである。

標準条件と fake good 条件は同程度に高い負の相関があることから、両条件において被検者は類似した回答を行っていたと考えられる。そのためにこの2条件間には抑うつ得点に差が出にくいかもしれない。また、D (NI<sub>30</sub>) 項目群の得点を見ると、標準条件において、項目の76~77% (13個中9.9~10.0個) が既に抑うつ採点キー側へ回答されている。D (NI<sub>50</sub>) 項目群でも71%

(22個中15.6～15.7個)が採点キー側へ回答されている。この天井効果のために FG 群においては fake good 教示によって得点の上昇する余地が少なく、逆説効果が生じなかったのではないかと考えられる。

## 第Ⅱ部 項目の回答に対する異常性の評定と reversal 現象

第Ⅱ部では、第Ⅰ部とは別の角度から精神的異常性の指標をとり、この指標を用いて、reversal 現象が見られるかどうかを確かめた。まず、項目に「あてはまる(はい)」もしくは「あてはまらない(いいえ)」と答えることについて、「全くおかしくない・正常である」から「非常におかしい・異常である」まで5段階に評定するよう評定者に求めた。次に、項目を抑うつ採点キー側回答に対する評定値が高い項目、即ち、より異常性の高い項目(以後これらの項目を便宜的に「D(AⅡ)項目」と記す)、評定値が低い、より正常性の高い項目(同じく「D(NⅡ)項目」と記す)に2分した。そして次に、標準手続き及び faking 教示の下での被検者の抑うつ得点を、D(AⅡ)項目群の得点とD(NⅡ)項目群の得点に分けて算出した。fake good 条件においては、D(AⅡ)項目群の得点が標準手続きよりも減少し、D(NⅡ)項目群の得点が上昇すれば reversal 現象が生じたといえる。また、fake bad 条件においては、D(AⅡ)項目群の得点が標準手続きよりも上昇し、D(NⅡ)項目群の得点が減少すれば reversal 現象が生じたといえる。D(AⅡ)項目とD(NⅡ)項目の分割点のとり方を2種類用いてみた。また、評定値とD側回答率の相関関係についても調べた。

## 方 法

### 被検者

**異常性の評定：**男子大学生311名を評定者とした。評定者は4種類の評定群の1つに割り当てられた。詳細は手続きの項に記した。

**faking：**fake good 条件群(FG 群)49名、fake bad 条件群(FB 群)34名、計83名の男子大学生を被検者とした。

実施はいずれも金沢大学教養部で開講している「心理学」の時間に行なった。

### 質問紙及び手続き

**異常性の評定：**MMPI(金沢大学版)の項目について、「あてはまる」(Y条件)、または「あてはまらない」

(N条件)と答えることが、どれくらい「精神的におかしい・異常だ」と思うかを5段階に評定するよう評定者に求めた(「全くおかしくない・正常である」を1、「どちらかといえばおかしくない・正常である」を2、「どちらとも言えない・わからない」を3、「どちらかといえばおかしい・異常である」を4、「非常におかしい・異常である」を5とした)。但し、MMPIの項目を前半(A条件)と後半(B条件)に二等分し、評定者にはどちらか一方について評定を求めた。即ち、評定者はYA、YB、NA、NBの4条件群のうちの1つに割り当てられた。割り当ては、4種類の小冊子を同数ずつ混ぜ込んで配布することによって行なった。各条件群に含まれた人数は、YA 群81名、YB 群77名、NA 群78名、NB 群75名となった。なお、ここでは抑うつ尺度に採点される項目のみを分析の対象にした。抑うつ尺度項目は、前半(A条件)に28項目、後半(B条件)に32項目が含まれていた。評定のペースは評定者に任せた。

**faking：**抑うつ尺度60項目とL尺度15項目、及びその他の項目2個を含む75項目(抑うつ尺度とL尺度で重複する項目が2個あるため)から成る質問紙を用いた。fake good 条件と fake bad 条件の教示は第Ⅰ部と同様であった。

まず全員に標準手続きで質問紙を実施し、次に fake good 条件と fake bad 条件のどちらか一方について回答を求めた。2種類の faking 用質問紙を同数ずつ混ぜ込んで配布し、被検者は自分が受け取った用紙に記入してある教示に従って、自己のペースで回答した。

## 結 果

### 抑うつ尺度項目の回答に対する異常性の評定

抑うつ尺度の各項目に対する異常性の評定の平均評定値とその標準偏差を表5に示した。また、平均評定値の高さからみた項目の分布を図1に示した。採点キー側回答については、評定値が3.00(どちらとも言えない)より大きい項目は22個、小さい項目は38個であった。評定値が4.00以上という異常性の高い項目は1個のみであり、2.00以下という正常性の高い項目は6個であった。60項目の評定値の総平均は2.70(SD=0.60)になり、正常側へやや偏っていた。なお、非採点キー側の総平均は2.50(SD=0.68)になり、両者に有意な差は見られなかった( $t=1.693$ ,  $df=59$ )。また、一つの項目で採点キー側と非採点キー側がどちらも3.00以上という項目はないが、どちらも3.00未満の項目は23個あった。「はい」と「いいえ」のどちらかに回答しても異常とは見なされない項目が $\frac{1}{3}$ 以上存在しているのであ

表5. 各項目の採点キー側, 非採点キー側に対する平均  
評定値及び標準偏差

項目番号 (原版)	採点キー側 平均評定値及び (標準偏差)	非採点キー側 平均評定値及び (標準偏差)
2	3.44 (1.039)	1.58 (0.739)
5	3.05 (1.123)	1.99 (0.893)
8	3.03 (1.013)	2.19 (1.014)
9	3.49 (0.935)	1.77 (0.902)
18	2.73 (1.408)	2.49 (1.304)
23	3.48 (1.050)	1.91 (1.047)
30	2.92 (1.236)	2.36 (0.991)
32	2.78 (1.012)	1.87 (0.858)
36	2.28 (1.115)	2.88 (1.088)
39	2.39 (1.184)	3.13 (1.092)
41	2.16 (0.812)	3.09 (0.982)
43	2.85 (1.152)	2.04 (1.025)
46	2.81 (1.135)	2.21 (0.853)
51	3.22 (1.147)	1.51 (0.691)
52	3.14 (0.956)	1.81 (0.881)
57	2.43 (1.016)	2.27 (1.072)
58	1.82 (1.041)	4.26 (0.853)
64	2.71 (0.955)	2.58 (0.978)
67	2.14 (1.073)	3.53 (1.319)
80	1.57 (0.857)	3.91 (0.846)
86	2.47 (1.062)	2.56 (0.988)
88	3.55 (1.180)	1.94 (0.885)
89	3.24 (1.130)	2.37 (0.980)
95	2.32 (1.016)	2.69 (0.977)
98	1.69 (0.822)	3.95 (0.887)
104	3.13 (0.937)	2.09 (0.857)
107	3.01 (1.059)	2.57 (1.152)
122	2.33 (1.057)	3.12 (0.932)
130	1.32 (0.637)	3.33 (1.455)
131	2.29 (1.088)	2.88 (1.013)
138	1.66 (0.736)	3.60 (1.151)
142	2.90 (1.071)	2.37 (1.136)
145	2.95 (1.161)	2.44 (1.107)
152	2.88 (1.112)	2.15 (1.074)
153	3.26 (1.171)	1.35 (0.602)
154	3.19 (1.216)	1.68 (0.799)
155	3.08 (1.171)	1.68 (0.910)
158	2.53 (1.152)	2.35 (1.017)
159	2.81 (1.038)	2.03 (0.911)
160	2.13 (1.024)	2.67 (1.045)
178	2.69 (1.048)	2.06 (0.927)
182	4.27 (0.822)	1.74 (1.098)
189	3.39 (0.891)	1.81 (0.800)
191	3.08 (1.136)	1.91 (0.764)
193	1.58 (0.772)	3.04 (1.232)
207	3.55 (1.031)	1.48 (0.681)
208	3.63 (1.152)	1.84 (0.873)
233	3.13 (0.963)	2.43 (0.768)
236	2.36 (0.899)	2.52 (0.940)
241	2.53 (1.131)	2.70 (1.113)
242	2.63 (1.010)	2.27 (0.853)
248	2.31 (1.085)	3.43 (0.954)
259	2.41 (1.070)	2.79 (1.024)
263	2.04 (1.189)	3.38 (1.179)
270	2.01 (0.987)	3.54 (0.949)
271	2.27 (1.040)	3.42 (0.947)
272	3.12 (0.929)	2.53 (1.021)
285	3.24 (1.071)	1.83 (0.818)
290	2.29 (0.958)	2.60 (1.103)
296	2.32 (0.918)	3.22 (1.210)

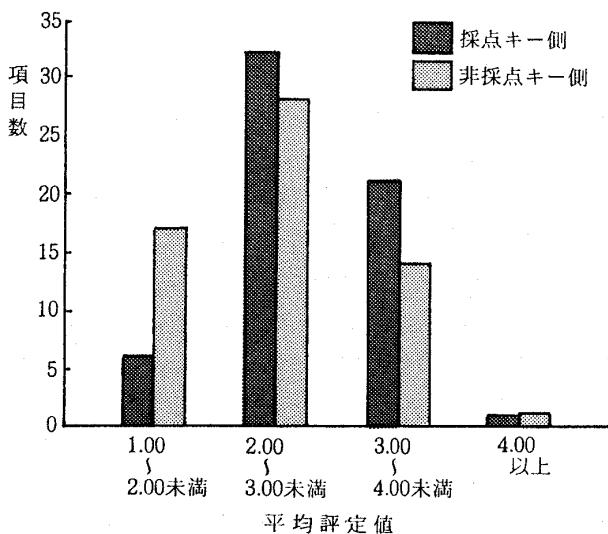


図1. 平均評定値の高さからみた項目分析

る。但し、採点キー側の評定値と非採点キー側の評定値の間には  $r = -0.860$  という負の相関が見られた。

faking 教示と reversal 現象

#### 1. 分割点を採点キー側 3.00 とした場合の reversal 現象の生起について

採点キー側評定値が 3.00 以上の22項目をD (A II) 項目, 残りの38項目をD (N II) 項目とみなした。前者をD (A II<sub>k</sub>) 項目, 後者をD (N II<sub>k</sub>) 項目と記す。各教示条件におけるD (A II<sub>k</sub>) 項目群とD (N II<sub>k</sub>) 項目群の抑うつ得点は表6と表7に示した。先ず, FG

表6. fake good 条件群におけるD(A II<sub>k</sub>) 項目群  
及びD(N II<sub>k</sub>) 項目群の抑うつ得点

	D (A II <sub>k</sub> ) 項目群	D (N II <sub>k</sub> ) 項目群	合計得点
標準手続き M	6.7	23.3	30.0 (T=52)
SD	3.08	3.56	5.31
fake good M	4.0	16.6	20.6 (T=36)
SD	1.73	2.78	3.61

(N=49)

採点キー側の平均評定値が3.00以上の22項目をD (A II<sub>k</sub>)項目, 3.00未満の38項目をD (N II<sub>k</sub>)項目と見なした場合の, D (A II<sub>k</sub>)項目群及びD (N II<sub>k</sub>)項目群の抑うつ得点を示した。

なお, ( )内のTの値は抑うつ尺度のT得点を示す。

表7. fake bad 条件群におけるD(A II<sub>k</sub>) 項目群  
及びD(N II<sub>k</sub>) 項目群の抑うつ得点

	D (A II <sub>k</sub> ) 項目群	D (N II <sub>k</sub> ) 項目群	合計得点
標準手続き M	6.2	22.4	28.6 (T=50)
SD	2.51	3.98	5.04
fake bad M	15.5	18.7	34.3 (T=59)
SD	3.67	4.15	7.07

(N=49)

採点キー側の平均評定値が3.00以上の22項目をD (A II<sub>k</sub>)項目, 3.00未満の38項目をD (N II<sub>k</sub>)項目と見なした場合の, D (A II<sub>k</sub>)項目群及びD (N II<sub>k</sub>)項目群の抑うつ得点を示した。

なお, ( )内のTの値は抑うつ尺度におけるT得点を示す。

群と FB 群の標準条件における抑うつ得点に既に差がないかどうかを確かめた。60項目全体から算出した抑うつ得点には FG 群と FB 群間に統計的な有意差は認められなかった ( $t = 1.205$ ,  $df = 81$ )。また, D (A II<sub>k</sub>) 項目群の得点 ( $t = 0.783$ ,  $df = 81$ ), D (N II<sub>k</sub>) 項目群の得点 ( $t = 1.081$ ,  $df = 81$ ) にも両

群間にそれぞれ有意な差はなく、FG 群と FB 群の被検者はほぼ等質であるといえる。

次に、reversal 現象の生起を調べた。FG 群においては、D (A II<sub>k</sub>) 項目群の得点 ( $t=6.071$ ,  $df=48$ ,  $p<0.001$ ), D (N II<sub>k</sub>) 項目群の得点 ( $t=11.624$ ,  $df=48$ ,  $p<0.001$ ) とともに標準条件より fake good 条件のほうが得点は低くなっていた。一方、FB 群においては、D (A II<sub>k</sub>) 項目群の得点 ( $t=-11.964$ ,  $df=33$ ,  $p<0.001$ ) は fake bad 条件のほうが得点が高くなっていたが、D (N II<sub>k</sub>) 項目群の得点 ( $t=3.464$ ,  $df=33$ ,  $p<0.001$ ) は低くなっていた。即ち、FB 群では reversal 現象が生起したが、FG 群では生起しなかった。但し、FG 群においても、D (A II<sub>k</sub>) 項目群の得点が標準条件よりも低くなるという現象は生起していた。つまり、reversal 現象が見出されなかったのは、D (N II<sub>k</sub>) 項目において逆説効果が生じなかったためである。

## 2. 非採点キー側評定値を考慮した場合の reversal 現象の生起について

上記は抑うつ採点キー側の評定値のみを用いた分割基準であった。上記の D (A II<sub>k</sub>) 項目、D (N II<sub>k</sub>) 項目の非採点キー側の評定値を見ると (表 5 参照), D (A II<sub>k</sub>) 項目の非採点キー側評定値はすべて 3.00 未満である。即ち、採点キー側が異常、非採点キー側が正常と評定されている項目である。しかし D (N II<sub>k</sub>) 項目は、①非採点キー側が 3.00 以上の項目、②採点キー側も非採点キー側も 3.00 未満であるが相対的には採点キー側の評定値の方が大きい項目、③同じく相対的に非採点キー側の評定値の方が大きい項目が含まれている。これらのうち、reversal 現象が最も起こりやすいのは①であろうと予想される。①の項目は採点キー側が正常、非採点キー側が異常と評定されているのに対し、②及び③は採点キー側も非採点キー側も正常と評定されているからである。そこで、①に含まれる15項目を新たに D

(N II<sub>nk</sub>) 項目と見なし、その他の45項目を D (A II<sub>nk</sub>) 項目と見なして抑うつ得点を算出した。結果は表 8 に示した。FG 群においては、D (A II<sub>nk</sub>) 項目群の得点 ( $t=11.902$ ,  $df=48$ ,  $p<0.001$ ) も D (N II<sub>nk</sub>) 項目群の得点 ( $t=2.146$ ,  $df=48$ ,  $p<0.05$ ) も標準条件より fake good 条件のほうが低かった ( $t=11.902$ ,  $df=48$ ,  $p<0.001$ )。つまり、FG 群においては reversal 現象は生起せず、それは D (N II<sub>nk</sub>) 項目において逆説効果が生じなかったためであった。なお、FB 群においては reversal 現象が生起していた (D (A II<sub>nk</sub>) 項目 ;  $t=-8.946$ ,  $df=33$ ,  $p<0.001$  ; D (N II<sub>nk</sub>) 項目 ;  $t=15.069$ ,  $df=33$ ,  $p<0.001$ )。この基準を用いた場合の各教示条件における D (N II<sub>nk</sub>) 項目群の得点分布を図 2 に示した。標準条件と fake good 条件の

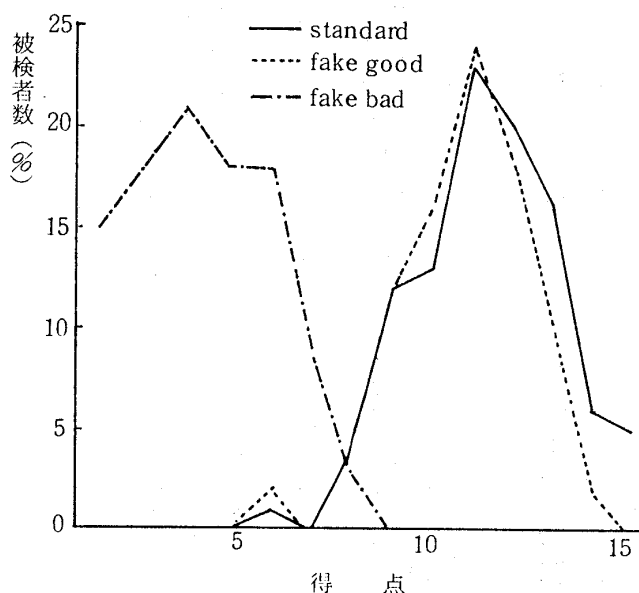


図2. D(N II<sub>nk</sub>)項目群の得点分布

分布がほぼ重なっていることと、どちらも高得点に偏り、fake good 教示によって得点の上昇する余地が少ないことがわかる。分割点を採点キー側評定値のみに基づいて決定すると D (N II<sub>k</sub>) 項目においては、採点キー側も非採点キー側も正常と評定される項目が含まれているために reversal 現象が生起せず、採点キー側が正常、非採点キー側が異常と評定される項目のみを D (N II<sub>nk</sub>) 項目と見なすと天井効果のために reversal 現象が生起しないのかもしれない。

結果として、D (A II) 項目と D (N II) 項目の分割点のとり方にかかわらず、FB 群では reversal 現象が生じたが FB 群では生じなかった。しかし、FG 群においても D (A II) 項目に対する回答は抑うつ得点が下

表8. fake good 条件群と fake bad 条件群における D(A II<sub>nk</sub>) 項目群及び D(N II<sub>nk</sub>) 項目群の抑うつ得点

		fake good条件群 (N=49)		fake bad 条件群 (N=34)	
		D(A II <sub>nk</sub> )	D(N II <sub>nk</sub> )	D(A II <sub>nk</sub> )	D(N II <sub>nk</sub> )
標準手続き	M	18.6	11.4	17.4	11.2
	SD	5.28	1.83	4.09	1.82
faking	M	9.9	10.7	29.9	4.4
	SD	2.97	1.54	6.80	1.67

採点キー側の平均評定値が3.00未満であり、且つ非採点キー側の平均評定値が3.00以上である15項目を D(N II<sub>nk</sub>)項目、それ以外の45項目を D(A II<sub>nk</sub>)項目と見なした場合の、各項目群の抑うつ得点を示した。

がる方向になされており、reversal 現象の説明のうち半分は成り立っていることを示した。これらは第Ⅰ部の結果と同じである。

### 3. 採点キー側評定値とD側回答率の相関について

項目の採点キー側評定値と各教示条件におけるD側回答率の相関を算出した。結果は表9のとおりであった。FG 群、FB 群ともに標準条件では割合高い負の相関が見られる（無相関検定：FG 群の標準条件  $t=-8.688$ ,  $df=58$ ,  $p<0.001$ ；FB 群の標準条件  $t=-9.650$ ,  $df=58$ ,  $p<0.001$ ）。一方 faking 教示下では、fake good 条件においては負の相関（無相関検定； $t=-5.651$ ,  $df=58$ ,  $p<0.001$ ）fake bad 条件においては正の相関（無相関検定； $t=5.068$ ,  $df=58$ ,  $p<0.001$ ）が見られる。標準条件及び fake good 条件では、異常

表9. 採点キー側平均評定値と各教示条件における採点キー側回答出現率の相関

	fake good 条件群	fake bad 条件群
標準手続き	-0.752	-0.785
faking	-0.532	-0.554

(N=60)

回答者は fake good 条件群49名、fake bad 条件群34名であった。

性評定値の高い（低い）項目ほど採点キー側に回答する人が少なく（多く）、fake bad 条件では評定値の高い（低い）項目ほど採点キー側に回答する人が多い（少ない）のである。これを各項目について考えると被検者は標準条件及び fake good 条件では項目の正常側に回答しており、fake bad 条件では項目の異常側に回答していたといえる。但し、相関の絶対値は標準条件よりも faking 条件のほうが低い。相関の符号が同じである FG 群について検定を行なったところ、fake good 条件のほうが標準条件よりも有意に低い相関を示した（ $t=-2.562$ ,  $df=57$ ,  $p<0.05$ ）。上記の、異常性評定値の高低と被検者の回答との関係は、faking 条件よりも標準条件においてより強いのである。

### 考 察

精神的異常性の指標を用いた場合には、reversal 現象は fake bad 条件では生じたが fake good 条件では生じなかった。fake good 条件においては逆説効果が生じなかったのである。この結果には、精神的異常性の指標と、標準条件におけるD側回答率との間に高い負の相関があることが関係していると思われる。即ち、被

検者は標準条件においても fake good 条件においても項目の正常側により多く回答しているため、両条件で項目の同じ側に回答することになり、得点に差が出にくいことが考えられる。特に採点キー側回答の異常性が低い項目では天井効果が生じ、fake good 教示によっても得点がそれ以上上昇する余地がなく、逆説効果が生じなかったと考えられる。ちなみに、項目の subtlety の程度と第Ⅰ部の各教示条件におけるD側回答率との相関は、FG 群の標準条件では  $r=-0.387$ 、fake good 条件では  $r=-0.727$ 、FB 群の標準条件では  $r=-0.390$ 、fake bad 条件では  $r=0.640$  であった（木村, 1987b）。問題率や異常性評定値との相関にくらべて標準条件の相関が低く、FG 群においても逆説効果が生じる余地があったと考えられる。事実、第Ⅰ部の被検者については subtlety 現象が認められており（木村, 1987b）、第Ⅱ部の被検者についても subtlety 現象が認められる（表10参照）。

表10. 第Ⅱ部の各教示条件における obvious (O) 項目群及び subtle (S) 項目群の抑うつ得点

		fake good 条件群 (N=49)		fake bad 条件群 (N=34)	
		O 項目群 <sup>1)</sup>	S 項目群 <sup>2)</sup>	O 項目群 <sup>3)</sup>	S 項目群 <sup>4)</sup>
標準手続き	M	20.9	9.1	19.9	8.7
	SD	6.00	1.83	4.86	1.59
faking	M	9.5	11.1	31.3	3.0
	SD	3.75	1.40	7.53	1.48

1)~4)の全てにおいて、標準手続きと faking 条件の得点の間には有意差が存在し、subtlety 現象が生じている。

- 1)  $t=12.882$ ,  $df=48$ ,  $p<0.001$     2)  $t=12.882$ ,  $df=48$ ,  $p<0.001$   
3)  $t=7.270$ ,  $df=33$ ,  $p<0.001$     4)  $t=14.116$ ,  $df=33$ ,  $p<0.001$

つまり、subtlety の指標を用いた場合は逆説効果が生じるのである。逆説効果は、採点キー側回答が精神的異常性を表わしているか否かといった、一般的な精神的異常性の指標によっては生起せず、どの回答が抑うつ尺度に採点されるかといった、項目内容から特定尺度を推測することで得られた指標によってのみ生起するといえる。

本研究では、reversal 現象の生起に関して第Ⅰ部と第Ⅱ部は同じ結果を示した。両者は異なった方法で「採点キー側へ回答することが精神的におかしいと判断される項目」と、「精神的におかしいとは判断されない項目」を分けている。前者にあたる項目は、第Ⅰ部におけるD (AⅠ) 項目、第Ⅱ部におけるD (AⅡ) 項目であり、後者にあたる項目は、第Ⅰ部におけるD (NⅠ) 項目、第Ⅱ部におけるD (NⅡ) 項目である。乱暴な言い方であるが前者を「異常」項目、後者を「正常」項目と呼ぶと、第Ⅰ部と第Ⅱ部における「異常」項目、「正常」項目はかなり重複しているのかもしれない。そこで

重複の度合いを調べてみた。第Ⅰ部で分割点を問題率50%にした場合と第Ⅱ部で分割点を採点キー側評定値3.00にした場合とは、「異常」項目21個、「正常」項目21個、計42個(70%)が重複していた。残り18項目のうち、1個は第Ⅰ部においてのみ「正常」とされ、17個は第Ⅱ部においてのみ「正常」とされた。第Ⅰ部で分割点を30%にした場合と第Ⅱ部で分割点を非採点キー側評定値3.00にした場合とは、「正常」項目11個、「異常」項目43個、計54個(90%)が重複していた。残り6項目のうち、第Ⅰ部においてのみ「正常」とされた項目が2個、第Ⅱ部においてのみ「正常」とされた項目が4個あった。項目の重複度は後の比較については非常に高く、両者はほぼ同じ項目群から成っていたといえよう。先の比較について重複度がやや低いのは、「異常」項目、「正常」項目の数がかもとも第Ⅰ部と第Ⅱ部においてかなり異なっているためである。最大限重複しても44個(73%)であるが実際にはそのうち42個と大部分が重複している。

第Ⅰ部と第Ⅱ部の結果における興味深い差異として、精神的異常性の指標として用いた測定値と、faking 条件におけるD側回答率との相関の違いが挙げられる。標準条件における相関係数の値には第Ⅰ部と第Ⅱ部の間に有意な差はない(FG群; $\chi^2=0.035$ ,  $df=1$ :FB群; $\chi^2=2.071$ ,  $df=1$ )が、faking 条件における相関には有意差が認められる(FG群; $\chi^2=46.216$ ,  $df=1$ ,  $p<0.001$ :FB群; $\chi^2=5.003$ ,  $df=1$ ,  $p<0.05$ )。第Ⅰ部における、問題率とD側回答率との相関のほうが、第Ⅱ部における、異常性評定値とD側回答率との相関よりも高いのである。この理由として考えられる要因を検討してみる。まず、各教示条件における第Ⅰ部と第Ⅱ部のD側回答率について相関を算出したところ、標準条件においては $r=0.914$ (FG群とFB群を込みにして算出した)、fake good 条件においては $r=0.976$ 、fake bad 条件においては $r=0.953$ となった。どれも高い正相関を示し、各教示条件のD側回答率には第Ⅰ部と第Ⅱ部で大きな違いはないといえる。第Ⅰ部における問題率と第Ⅱ部における採点キー側評定値との相関は $r=0.737$ である。両者はある程度まで共通のものを測っているが、全く同じものを測っているわけでないことを示している。第Ⅰ部の被検者は女性であり、第Ⅱ部の被検者は男性であることから、第Ⅰ部と第Ⅱ部における精神的異常性の指標とD側回答率の相関の違いには性差が影響していることも考えられるが性差のみで説明されるかどうかは疑問である。faking 回答の被検者も同じように女性と男性に分かれているが両者の回答の相関は0.9以上と高いからである。第Ⅰ部は「はい」か「いいえ」を強制的

に選ばせる形で異常性の指標を求めており、第Ⅱ部はどちらか一方の回答のみについて異常性の指標を求めているということの違いかもしれない。ただし、この違いの影響だけでは、標準条件において両者が同じ程度の相関を示していることの説明はしにくいと思われる。ほかに、「精神的におかしい・問題がある」という教示と「精神的におかしい・異常である」という教示の違いによる影響も考えられる。faking 教示に従って回答する時に被検者が基準としているであろう次元が、それがどんなものであれ、「異常である」という言葉の持つニュアンスよりも、「問題がある」という言葉の持つニュアンスと共通性が大きいのかも知れない。

reversal 現象が見出されている他の研究には、Wales and Seeman (1968)の、X項目—ゼロ項目の分類を用いた fake good 条件における研究、Harvey and Sipprelle (1976)による、Wienerの subtlety 分類を用いた研究、Burkhart, Christian, and Gynther (1978)による、独自の subtlety 指標(Christian, Burkhart, and Gynther; 1978)を用いた研究などがある。ゼロ項目、subtle 項目がそれぞれ逆説効果を生じているのである(ゼロ項目とは正常被検者による多数側回答が臨床尺度の採点キー側回答になっている項目であり、X項目とは正常被検者の少数側回答が臨床尺度の採点キー側回答になっている項目である)。このうち Burkhart et al. (1978)は、ある項目がどの臨床尺度に採点されるかということは無視して、「心理的問題点をはっきり表わしている程度」を被検者に very subtle から very obvious まで5段階に評定させて subtlety の指標としている。本研究と同じように、特定の尺度と項目の関連ではなく、より一般的な精神的異常性について尋ねる方法で指標を求めながらも、Burkhart et al. (1978)は fake good 条件においても fake bad 条件においても reversal 現象を見出している。但し、彼らの研究では MMPI 全尺度の中での subtle 項目群の合計得点の変化を調べており、今回の研究のように1つの尺度内における reversal 現象の生起を調べてはいない。これは他の2研究についても同様である。本研究で用いた精神的異常性の指標についても全尺度を通した合計得点の変化を求めれば fake good 条件、fake bad 条件共に reversal 現象が認められるかもしれない。

## 要 約

MMPI 抑うつ尺度項目を、採点キー側回答が推測しやすい obvious 項目と、推測しにくい subtle 項目



に分類すると、faking 教示下では obvious 項目は被検者の意図する側へ、即ち fake good 条件においては抑うつ得点が低くなる方向へ、fake bad 条件においては抑うつ得点が高くなる方向へ回答が変化した。しかし、subtle 項目は被検者の意図する側とは反対側へ、即ち fake good 条件においては抑うつ得点が高くなる方向へ、fake bad 条件においては抑うつ得点が低くなる方向へ回答が変化した（木村，1987b）。本研究はこれと同等の現象が、木村（1987b）で用いたような subtlety の指標ではなく、項目の一般的な精神的異常性を表わす指標を用いた場合にも認められるか否かを検討した。この現象を reversal 現象と呼び、reversal 現象の一部である「精神的におかしいと判断されない項目においては意図とは逆の方向に回答する」ことを、Burkhart, Christian, and Gynther (1978) に倣って「逆説効果」と呼んだ。精神的異常性の指標として2つの指標を用いた。一つは「はい」、「いいえ」のうち、どちらに答えると「精神的におかしい・問題がある」かの判断に基づく指標である。もう一つは、項目に「あてはまる」または「あてはまらない」と答えることが「どれくらい精神的におかしい・異常であるか」の評定に基づく指標である。その結果、どちらの指標を用いた場合も、reversal 現象は fake bad 条件においては生じたが fake good 条件においては生じなかった。これは逆説効果が生じていないためであった。精神的異常性の測定値と標準教示における抑うつ採点キー側回答率との間に高い負の相関があり、異常性が低い項目については標準手続きにおいて既に高率で採点キー側へ回答されているので、fake good 教示によって、より以上に高い得点をとる余地がないことが原因ではないかと考えられた。

## 参 考 文 献

- Burkhart, B.R., Christian, W.L., and Gynther, M.D. 1978 Item subtlety and faking on the MMPI: A paradoxical relationship. *Journal of Personality Assessment*, 42, 76-80.
- Christian, W.L., Burkhart, B.R., and Gynther, M.D. 1978 Subtle-obvious ratings of MMPI items: New interest in an old concept. *Journal of Consulting and Clinical Psychology*, 46, 1178-1186.
- Harvey, M.A. and Sipprelle, C.N. 1976 Demand characteristic effects on the subtle and obvious subscales of the MMPI. *Journal of Personality Assessment*, 40, 539-544.
- 木村敦子 1987a MMPI における抑うつ尺度項目の subtlety 測定の試み 金沢大学文学部論集 行動科学科編, 8, 23-34.
- 木村敦子 1987b MMPI D尺度における faking の効果と項目の subtlety との関係 臨床心理学の諸領域—金沢大学臨床心理学研究室紀要—, 6, 39-45.
- 多田治夫 1959 MMPI 邦語版の作成と大学生群の結果 金沢大学法文学部論集 哲学史学篇, 7, 137-172.
- 田中富士夫 1964 MMPI 邦語版標準化の試み（中間報告）—MMPI 金大版の改定とその資料— 金沢大学法文学部論集 哲学篇, 12, 71-97.
- Wales, B. and Seeman, W. 1968 A new method for detecting the fake-good responses set on the MMPI. *Journal of Clinical Psychology*, 24, 211-216.

# Relationship between Paradoxical Effect under Faking Condition and Abnormality on the Items of MMPI D Scale

Atsuko Kimura

Kanazawa University

When subjects are asked to fake on the MMPI, they are able to manipulate their responses to the obvious items successfully, but not to the subtle items. Furthermore, on subtle items, they responded in the opposite direction to their intention ; Subjects obtained higher scores on the scale D under faking-good condition, and lower scores under faking-bad condition (Kimura, 1987b). Burkhart et al. (1978) called this phenomenon as paradoxical effects.

The present study tries to determine if the paradoxical effects have any relation with item abnormality. Two kinds of abnormality measures were used. The first one was obtained by

asking subjects "which response of true or false is more abnormal". The second measure was obtained by asking "to what extent is this response abnormal".

The paradoxical effects were demonstrated under faking-bad condition. Under faking-good condition, however, the effects were not observed. This finding may be explained by item abnormality in terms of both measures described above. For the less-abnormal items, our subjects could not get higher scores, even if it was faking-good condition, than the high scores obtained under standard condition.